

VU Research Portal

An application of the Tree-Step Test-Interview (TSTI): A validation study of the Dutch and Norwegian versions of the 'Illegal Aliens Scale'

Hak, T.; van der Veer, C.G.; Ommundsen, R.

published in

International Journal of Social Research Methodology
2006

DOI (link to publisher)

[10.1080/13645570600656264](https://doi.org/10.1080/13645570600656264)

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Hak, T., van der Veer, C. G., & Ommundsen, R. (2006). An application of the Tree-Step Test-Interview (TSTI): A validation study of the Dutch and Norwegian versions of the 'Illegal Aliens Scale'. *International Journal of Social Research Methodology*, 8(Juli 2006), 215-227. <https://doi.org/10.1080/13645570600656264>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

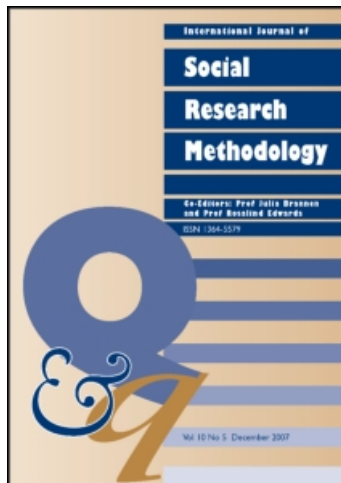
This article was downloaded by: [Vrije Universiteit, Library]

On: 26 May 2011

Access details: Access Details: [subscription number 907218003]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Social Research Methodology

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713737293>

An Application of the Three-Step Test-Interview (TSTI): A Validation Study of the Dutch and Norwegian Versions of the 'Illegal Aliens Scale'

Tony Hak; Kees van der Veer; Reidar Ommundsen

To cite this Article Hak, Tony , van der Veer, Kees and Ommundsen, Reidar(2006) 'An Application of the Three-Step Test-Interview (TSTI): A Validation Study of the Dutch and Norwegian Versions of the 'Illegal Aliens Scale'', International Journal of Social Research Methodology, 9: 3, 215 – 227

To link to this Article: DOI: 10.1080/13645570600656264

URL: <http://dx.doi.org/10.1080/13645570600656264>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

An Application of the Three-Step Test-Interview (TSTI): A Validation Study of the Dutch and Norwegian Versions of the 'Illegal Aliens Scale'

Tony Hak, Kees van der Veer & Reidar Ommundsen

Received 12 January 2004; Accepted 4 March 2005

In this paper the results of a validation study of the Dutch and Norwegian versions of the Illegal Aliens (IA) Scale are discussed. The main objective is to evaluate the usefulness of the Three-Step Test-Interview (TSTI) as a pre-testing method in a validation study. The TSTI is a new technique specifically developed to detect problems with self-completion questionnaires. The study's aims were twofold: (1) to evaluate the usefulness of the TSTI for detecting problems regarding attitude measurement; and (2) to identify potential problems with the IA Scale. The IA Scale is a Likert-type attitude scale consisting of 20 parallel interval items. Its main purpose is for large comparative sample studies of political and ideological attitudes, for example between subgroups within populations or between countries. The TSTI procedure appeared to be very productive in terms of the number and kind of problems that were found regarding the Dutch and Norwegian versions of the IA Scale. It is concluded that the TSTI is useful as a diagnostic tool in validation studies of instruments that are aimed at measuring attitudes.

Introduction

In this article we discuss results of a part of a validation study of the Dutch and Norwegian versions of the Illegal Aliens (IA) Scale (Ommundsen & Larsen, 1997). This

Tony Hak is at the Department of Management of Technology and Innovation, RSM Erasmus University, Rotterdam, The Netherlands. Kees van der Veer is at the Department of Social Science Methodology, Vrije University, Amsterdam, The Netherlands. Reidar Ommundsen is at the Institute of Psychology, University of Oslo, Oslo, Norway. Correspondence to: Tony Hak, Department of Management of Technology and Innovation, RSM Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands. Tel.: +31 10 4089594; Fax: +31 10 4089016; Email: thak@rsm.nl

research is part of an international project with participants in the United States (US), Norway, Denmark, Spain, Bulgaria, Vietnam, Australia, New Zealand and the Netherlands (van der Veer, Ommundsen, & Hak, 2004). The Three-Step Test-Interview (TSTI) (Hak, van der Veer, & Jansen, 2004) is a qualitative research instrument specifically designed for testing self-completion questionnaires. It was developed and tested in a study of the quality of alcohol consumption questions (Jansen & Hak, 2005). That study suggested that the TSTI might be particularly good at identifying problems that originate from a mismatch between, on the one hand, the 'theory' underlying the questions and, on the other hand, features of a respondent's actual behaviour and biography. These problems were not identified by an *expert review* that was conducted alongside it. Arguably, they will also not be identified by 'cognitive' techniques because the usual methods (such as probing about definitions of concepts and paraphrasing tasks) do not address these issues. In this article we applied the TSTI method for the first time to an attitude scale.

Responding to an attitude scale is a cognitive task that is very different from answering factual questions about one's alcohol consumption in the last week or month. The first aim of the present study was to find out whether the TSTI would be equally productive in discovering problems with an attitude scale as it had been with the alcohol consumption questions, and to see which kind of problems would be found. The second aim was to identify potential problems with the IA scale.

This paper begins with a discussion of the background, assumptions and procedures of the TSTI. The IA Scale is then described. The main part of this article discusses the results of the application of the TSTI to the Dutch and Norwegian versions of the IA scale. Finally, conclusions are drawn, both about the usefulness of the TSTI for testing attitude questionnaires and about the IA scale.

The Three-Step Test-Interview (TSTI)

The TSTI is in many respects similar to the *cognitive interview* used in survey institutions such as the Census Bureau and the Bureau of Labor Statistics for pre-testing questions (see Willis, 2005). The main similarity between the TSTI and the cognitive interview is the use of *think aloud* techniques (Ericsson & Simon, 1984, 1993). The term *cognitive interview* was found to be confusing because, in the psychological literature, this term is usually applied to a technique of investigative interviewing used by the police (see e.g. Bekerian & Dennett, 1993). In order to avoid confusion between these two very different techniques of data collection and of validating evidence, we prefer the more neutral term *qualitative test-interview* as a generic term for various techniques that are used to evaluate the quality of questionnaires. Whereas qualitative test-interviews are mainly used in the *pre-testing* of survey questionnaires, the TSTI could be used to test the quality of survey data throughout the survey process, not only in the pre-testing stage.

In the literature on (pre-)testing, two main cognitive interview techniques are usually discussed: *think aloud* and *probing* (see e.g. Willis, 2005). These 'cognitive' techniques are used to assess respondents' difficulties regarding issues such as the

comprehension of the task, the interpretation of concepts and definitions, the complexity of the task, the understanding of questions, and the ease of recording answers. In the literature and in published manuals there is a blurring of three very different objects of cognitive (pre-)testing:

- (1) *actual* difficulties encountered by respondents when they complete a questionnaire;
- (2) *potential* difficulties;
- (3) retrospective *accounts* of what happened during the completion of the questionnaire.

Published lists of recommended *probes* do not distinguish between these three objects of (pre-)testing. Examples of such probes are:

- *concurrent think aloud*: 'As you answer the questions, please try to say in words what you are thinking'
- *retrospective think aloud*: 'Please try to say in words, in as much detail as possible, what you did when you answered the question'
- *comprehension probe*: 'Could you tell me in your own words what this question is asking?'
- *recall probe*: 'How did you remember that?'
- *confidence judgment probe*: 'How well do you remember this?'
- *response probe*: 'How did you feel about answering this question?'

In our view it is important to distinguish between studying *actual* response behaviour, *accounts* of actual behaviour and *potential* behaviour. Generally, survey researchers and (pre-)testers have been interested too much in *accounts* of past behaviour and in *potential* difficulties that respondents might encounter. The focus should be more on *actual response behaviour*. The TSTI has been developed specifically as an instrument for discovering problems that occur during the completion of self-administered questionnaires by *observing* actual response behaviour.

Because much of actual response behaviour consists of 'thinking' and is therefore hidden from the observer, the TSTI uses the technique of (*concurrent*) *think aloud* to make this behaviour 'observable'. Think aloud is a difficult technique for the respondent and often yields only partial data. It should ideally be followed by *retrospective think aloud*, in order to fill in the gaps. According to Ericsson and Simon (1993), retrospective think aloud must take place within five minutes of completing the task. Therefore this technique can only be correctly applied to individual items or to very short scales or questionnaires. This implies that the technique of retrospective think aloud cannot be used for a study of how respondents actually behave when they complete a (longer) scale or questionnaire. Gaps of concurrent think aloud can still be filled in by means of retrospective *probing*, but the interviewer (tester) must be aware that (and in what respects) retrospective probing is different from retrospective think aloud. In order to avoid any misunderstanding, we refer to retrospective probing (as practised in the TSTI) as 'focused interviewing'.

The aim of the TSTI is to collect data on how respondents actually complete a questionnaire. The aim is explicitly *not* to test individual items separately from their actual context in a questionnaire. The TSTI consists of the following steps.

Step 1. Concurrent Think Aloud Aimed at Collecting Observational Data

This step produces two types of data:

- (1) Observations of respondent behaviour (such as skipping questions, correction of the chosen response category, hesitation, distress etc.).
- (2) Think aloud data. Respondents are instructed to complete the questionnaire as they normally would do, with the additional task of concurrently verbalizing what they are thinking.

Ideally, both types of observational data—actions and verbalizations—are recorded and kept on audiotape and/or videotape for later analysis. But the interviewer also makes ‘real-time’ notes of observed behaviours and of verbalized thoughts that seem to be indicative of problems in the response process. These real-time notes are made for immediate use in the following steps of the interview.

The strictly observational nature of this first and essential step of the TSTI must not be compromised by any intervention—such as a question, comment, probe—by the interviewer that might suggest that a self-report from the respondent is required. Respondents must be asked—and perhaps be trained—to only report the actual stages of their thinking during the completion of the task, i.e. when responding to the questions of the questionnaire. They must be asked explicitly *not* to explain or justify their thoughts to the interviewer.

Step 2. Focused Interview Aimed at Clarifying and Completing Observational Data

In this step, interviewers only consider those actions or thoughts of the respondent which they have observed (in step one) and do not feel fully informed about. This step will enable him to fill in the gaps in the observational data or to check information (e.g. ‘Did I hear you say ...?’ or ‘You stopped for a while there, what did you think?’). The assessment of this (in)completeness must be made in real time based on the interviewer’s observation notes. The main methodological criterion (and also technically the most difficult aspect of this step for both respondents and interviewers) is that respondents should only report about what they did and thought in the first step, *not* about what they think now (in retrospect). It is *not* the aim of this step to elicit accounts, comments etc.

Step 3. Semi-structured Interview Aimed at Eliciting Experiences and Opinions

This final step is the only one in the TSTI in which the respondent is ‘allowed’ and even stimulated to add secondary data to the primary, observational ones. Examples of such secondary data are accounts and reports of feelings, explanations and preferences. This step may take very different forms depending on the kind of questionnaire that is being ‘tested’. Three main forms (and corresponding aims) can be distinguished, as follows.

- (a) Respondents might (be requested to) ‘explain’ their response behaviour—particularly when specific problems were encountered in responding to the questionnaire.

They could comment on what they thought the exact nature of the problem was and why they behaved as they did—as recorded in steps 1 and 2 of the ‘interview’. Also, respondents might suggest improvements in terms of wording of questions, layout of the questionnaire, instructions etc. The aims and form of this interview will be similar to those of ‘respondent debriefing’. It is important to be aware that these kinds of comments constitute ‘opinions’, not facts, regarding the causes of problems detected in steps 1 and 2. Interviewers must make their own analysis of problems associated with the questionnaire (based on observations in all interviews of the (pre-)test).

- (b) Respondents might be asked to paraphrase questions and to comment on their definitions of terms. In other words, some form of ‘cognitive interviewing’ (probing) might be done in this stage of the TSTI.
- (c) Respondents might be asked to give additional information about the substantive issues that are covered by the tested questionnaire. For instance, if an alcohol consumption questionnaire is tested, respondents might be invited to describe their alcohol consumption in their own words. Or, if a scale for the measurement of attitudes towards ‘illegal immigrants’ is tested, respondents might be asked to explain these attitudes in their own words. In our first pilot study concerning alcohol consumption (Jansen & Hak, 2005), it appeared that such data from this part of the interview, when compared to respondents’ responses to the questionnaire in step 1, were useful as indicators of the validity and reliability of the data collected by the questionnaire.

The forms (a) and (b) of this third step may seem very similar to usual formats of ‘cognitive interviewing’. It must, however, be emphasized that data collected in step 3 of the TSTI have a very different status from those usually collected in cognitive interviews. The main difference is that they are collected additionally (and secondarily) to primary, observational data. They are elicited as aids in the analysis of the primary data on *actual* response behaviour (and on actual problems that occur) rather than as primary data about (by definition) *potential* problems. It should be noted, therefore, that in the strict sense of the term, the TSTI is *not* an interview. Rather, it is a sequence of (a) observation, (b) follow-up probing and (c) validation.

The TSTI is primarily an *observation* instrument. Strict adherence to the three steps as described, in this order, should guarantee that the *actual* behaviour of the respondent is observed (step 1) and reconstructed (step 2) before any interviewing (step 3) occurs.

A Validation Study of the Illegal Aliens (IA) Scale

Since the 1970s, the issue of illegal immigration has become a matter of concern to people in many parts of the world. Economic dislocation, ethnic strife in Eastern Europe and Africa, and fundamentalism in religion and politics result in increasing numbers of refugees. Since options for legal immigration are limited, many seek

security and a better life illegally in developed nations. This phenomenon of illegal immigration evokes negative attitudes towards those who are considered 'illegal' immigrants. Since it usually is unclear which immigrant is to be considered 'legal' and which one 'illegal', negative attitudes are often directed towards the broad and vague category of immigrants, including 'legal' political refugees and asylum seekers. Given this context in which the issue of illegal immigration has become more and more salient in the Western world, it is relevant to find out how people in different countries respond to this immigration situation and what attitudes they express towards 'illegal aliens'.

Within this context, Ommundsen & Larsen (1997) have developed a Likert-type attitude scale, the Illegal Aliens (IA) Scale, consisting of the following 20 parallel interval items.¹

- (1) Illegal aliens should not benefit from my tax money.
- (2) Our taxes should be used to help those residing illegally in Norway/the Netherlands.
- (3) There is enough room in this country for everyone.
- (4) Illegal aliens are not infringing on our country's resources.
- (5) Illegal aliens are a nuisance to society.
- (6) There should be open international borders.
- (7) Access to this country is too easy.
- (8) Illegal aliens should be excluded from social welfare.
- (9) Norway/the Netherlands should accept all political refugees.
- (10) Illegal aliens who give birth to children in Norway/the Netherlands should be made citizens.
- (11) Illegal aliens cost Norway/the Netherlands millions of kroner/guilders/euros each year.
- (12) Illegal aliens should be eligible for welfare.
- (13) Illegal aliens provide Norway/the Netherlands with a valuable human resource.
- (14) The government should pay for the care and education of illegal aliens.
- (15) Illegal aliens should not have the same rights as Norwegian/Dutch citizens.
- (16) Illegal aliens have rights too.
- (17) Taking care of people from other nations is not the responsibility of Norway/the Netherlands.
- (18) All illegal aliens deserve the same rights as Norwegian/Dutch citizens.
- (19) Illegal aliens should be forced to go back to their own countries.
- (20) Illegal aliens should not be discriminated against.

Each of these statements is followed by five response options:

Agree strongly	1
Agree	2
Uncertain	3
Disagree	4
Disagree strongly	5

The main purpose of the IA Scale is to be used in large comparative sample studies of political and ideological attitudes, such as comparing subgroups within populations or between countries. For cross-national research purposes, the IA Scale has been translated into Norwegian, Danish and Dutch. The Norwegian and Dutch versions of the scale have been subjected to a series of validation studies (see Ommundsen, van der veer, Mørch, Hak, & Larsen 2002; van der Veer *et al.*, 2004).

This study is part of a series of studies. The specific aims of this part were, first, to describe the range of possible interpretations of the statements of the IA Scale by Norwegian and Dutch respondents and, second, to explore possible reasons for different interpretations. Preceding the TSTI study, an *expert review* was conducted in which experts in survey research identified problems regarding the wording of a number of statements. This expert review was conducted to find out whether the TSTI would identify a smaller or larger number of problems than the expert review and whether problems identified with the TSTI would be the same or different from the problems identified by experts. For the TSTI study, two convenience samples were recruited, consisting of six and eight undergraduate students in the social sciences at the Vrije University Amsterdam and the University of Oslo respectively.

The TSTI study identified more or less the same wording problems as were identified by the expert review. These results will not be presented here. Instead we will focus on the TSTI results that pertain to interpretations of the items of the IA Scale in the two populations. These problems regarding the interpretation of items were not identified by the expert review. When we present results below, we will quote quite extensively from the verbal *protocols* (transcripts) of the TSTI sessions in order to document how results were produced in the interviews.

Results and Discussion

Our main finding (both in the Norwegian and the Dutch sample) regarding the interpretation of items of the IA Scale was that a number of respondents felt that in a number of cases they must make a choice between two possible 'readings'. Take the following example.

Respondent 5 NL, Item 11

Illegal Aliens Cost the Netherlands Millions of Guilders Each Year

Step 1. Concurrent think aloud

- R: *illegal aliens cost the Netherlands millions of guilders each year ...in first instance I don't know anything about it, so I must make an intuitive choice ... and I think that this is the case, but I don't say that it's negative and then I don't know how I must interpret this item, because I would ... with such an item I would rather, ... well, think perhaps in a socially desirable way, whereas, as such, the item ... the item does not express an opinion, so I choose for just the literal, interpretation ... and that, yes, it's a fact that it costs millions of, millions of guilders, so I agree [response category 2]*

Step 2. Focused interviewing

- R: yeah ... here I immediately thought ... yes that's correct ... but ... I don't know whether I have said that ... but what I intuitively feel a bit ... is that if you ... if you *agree* strongly with it, it looks like you are very much anti-illegal aliens
- I: yes
- R: but the statement as such does not say anything about attitudes to illegal aliens ... and then I think it's their fault that they don't have formulated it differently ... I think I should respond to the statement that is actually there
- I: okay ... and why have you
- R: chosen the 2 (agree) rather than 1 (agree strongly)?
- I: yes ... because you are convinced that illegal aliens cost us millions of guilders
- R: well, I think again that I have chosen 2 because, well, I am almost certain that it costs so much but ... well on what base?
- I: oh ... so it's the uncertainty ... that you do not know the facts
- R: yes

The respondent explicitly notices two possible interpretations of this item. One is the 'literal' one: do illegal aliens cost The Netherlands millions of guilders each year, or don't they? The respondent thinks that they do. But he is aware of a possible other interpretation than the 'literal' or 'factual' one, namely a (let us say) 'political' one in which the item does not measure the respondent's idea about the economic facts but rather a hostile or friendly attitude to illegal immigrants. Let us look at another example from the same TSTI session.

*Respondent 5 NL, Item 13**Illegal Aliens Provide the Netherlands with a Valuable Human Resource**Step 1. Concurrent think aloud*

- R: *illegal aliens provide the Netherlands with a valuable human resource* ... I immediately think, in my opinion it doesn't make a difference whether you are legal or illegal, to be a valuable human resource, so, well, I have not ... a straightforward opinion ... so, it is uncertain [response category 3] ... because one can be valuable also if one is legal.

Step 2. Focused interviewing

- R: Well ... here I have ... I think, in my opinion, I lack knowledge a bit ... I have *uncertain* ... well, *a valuable human resource*, well, what can I say about it?
- I: well there might also be a kind of logical reasoning behind it ... that the item suggests a difference between illegal and legal people
- R: oh right yes that's what I said
- I: you said if this applies to everyone, why should I confirm it here for illegal people only? that kind of reasoning
- R: yes yes could be yes
- I: but now ... you say, you say also the facts ... you don't know the facts
- R: no, you're right, the way I thought was ... indeed I thought that the distinction between illegal or normal, just a normal dutch person ... that is not clear to me ... and if there is a distinction ... well, I don't know either whether they are valuable or damaging ... those are two things, a bit two things ... of which I don't know

In the transcript of the concurrent think aloud step it can be observed that the respondent recognizes the item as one in which a difference is constructed between 'legal' and 'illegal' immigrants. He disagrees with this distinction. His selection of the response category *uncertain* can be seen as expressing an avoidance of taking sides for or against 'illegal' immigrants.

In TSTI step 2 (focused interviewing), the respondent gives another reason for choosing *uncertain*, namely that he is genuinely uncertain whether illegal aliens actually are a valuable human resource in the economy. In this reasoning, the respondent interprets the statement 'literally', not as the expression of a hostile or friendly attitude towards immigrants.

In the third step of the TSTI (semi-structured interviewing; no extract due to length considerations), this respondent confirmed that he was aware of the fact that the resulting IA score was less 'friendly' towards illegal immigrants than would have been the case if he had followed the expectations of the authors of the questionnaire. He had clearly recognized that the authors would expect him to demonstrate his friendly attitude to illegal immigrants wherever possible, i.e. by reading items as invitations to position himself politically rather than as questions about economic facts. He described himself as someone who tends to 'interpret everything always very literally'. This self-description explains how the wording of items has made it possible for this respondent to find a lack of clarity in many items and to justify a 'literal' reading of these.

At the end of the third TSTI step, the respondent's strategy and its implications were discussed explicitly with him (see the next extract).

Respondent 5 NL, Item 11

Illegal Aliens Cost the Netherlands Millions of Guilders Each Year

In step 1 (think aloud) this respondent had answered *agree* [response category 2].

Step 3. Semi-structured interviewing

- I: take the item *illegal aliens cost the Netherlands millions of guilders each year* ... such an item ... you take it literally
- R: erm
- I: you you know for sure that a couple of millions is not much so it's very likely that illegal aliens cost us millions
- R: yes
- I: do you think ... so you interpret the item as a statement about facts but is it right that you assume that the designers of the questionnaire have something else in mind?
- R: yes
- I: now if this questionnaire is a test in logic ... you have performed very well on the test
- R: yes (laughs)
- I: but if it is true that ... if this questionnaire aims at measuring ... say your benevolence regarding illegal aliens ... in that case you have been almost deliberately, deliberately
- R: yes on the wrong side
- I: so you mislead the researchers ... so one can say that for that reason alone the item does not measure your ... what your real opinion is about illegal aliens

- R: no ... not at all
 I: and you say ... well it is their responsibility to ... how they interpret the responses ... now it is possible that ... if they want to draw political conclusions ... that they as you say will interpret your response incorrectly
 R: yes ... but I try to attend to what ... as much as possible ... to what is printed here ... that's in principle the only thing I have

At this point we would like to draw two conclusions, one concerning the IA Scale and the other concerning the TSTI. Our conclusion regarding the IA Scale is that there *might* be a problem with the scale in the sense that 'friendly' attitudes to immigrants might be under-represented in IA Scale results. This possibility deserves further research. Our conclusion regarding the TSTI is that it appears to be very productive in detecting this kind of problem that might compromise measurements with the IA Scale. However, so far we have only discussed one example from the Dutch sample. We will now expand upon this first result by looking at examples from the Norwegian sample.

The first example is with regard to Item 11 again.

Respondent 2 NOR, Item 11

Illegal Aliens Cost Norway Millions of Kroner Each Year

Step 1. Concurrent think aloud

- R: I don't have the slightest idea ... I assume they cost some money for the Government to check them out ... But as I say I don't know. I'll say 3 [uncertain]

Step 2. Focused interviewing

- R: What does it mean ... I think of Carl I. Hagen.

The similarity with the Dutch example, discussed above, is striking. This respondent also notices that two interpretations of the item are possible, a literal one for which the appropriate response category would be *uncertain* ('I don't have the slightest idea') and a political one, for or against Carl I. Hagen, the leader of the chauvinistic Fremskrittspartiet (Progress Party). A remarkable difference between the Dutch and Norwegian respondents is that the latter mention often and more explicitly that they recognize certain items as belonging to a discourse that is 'hostile' to immigrants.

Sometimes Carl I. Hagen was explicitly mentioned as a potential author of a statement of the scale, as in the two following examples.

Respondent 2 NOR, Item 1

Illegal Aliens Should Not Benefit from My Tax Money

Step 1. Concurrent think aloud

- R: I have some problems with the term illegal aliens ... Apart from that it's pretty clear. It's something about the way you say it though. It's like what you hear in relation to

immigration. Opponents would talk about 'enjoy tax money'. It sounds a bit like Carl I. Hagen. [response category 4, disagree]

Respondent 3 NOR, Item 1

Illegal Aliens Should Not Benefit from My Tax Money

Step 1. Concurrent think aloud

R: Here it's very much MY tax money, it's all about person. So what do I mean by this ... I don't want them to enjoy my taxes, but on the other hand ... But in a way they're coming from some place that is not good. I can't say I totally agree. It's difficult, I do have an opinion, but ... I think I'll say disagree. If I had been Carl I. Hagen I would totally agree. I think I *disagree*. [response category 4]

The recognition of some items as belonging to the repertoire of Carl I. Hagen implies that respondents must decide between two ways of responding. One way consists of taking a stance vis-à-vis Hagen's discourse, which is hostile to immigrants. The other consists of sticking to a 'literal' or 'factual' interpretation of the item.

Other respondents do not mention Carl I. Hagen in particular, but relate some items of the scale to chauvinist and populist opinions (which they detest) in general, such as in the next two examples in which respondents refer to the 'tabloids' and to a 'fascistic granduncle'.

Respondent 4 NOR, Item 1

Illegal Aliens Should Not Benefit from My Tax Money

Step 1. Concurrent think aloud

R: Basically I think it's a reasonable statement, because if they get too much of our taxes it would stimulate to more illegal immigration ... On the other hand it is a normative statement, and I regard it as a question of principles ... And then of course, it's not right, because it will always be someone with a true need for at least some of my tax money, whether they are illegal or not. So I *disagree*. [response category 4]

Step 2. Focused interviewing

R: When I see this I think about tabloids ...

I: What do you mean?

R: 'Here they come and get your hard earned money'.

Respondent 6 NOR, Item 1

Illegal Aliens Should Not Benefit from My Tax Money

Step 1. Concurrent think aloud

R: Well, on this one I think ... I picture my fascistic granduncle in Sjøk who is the prototype of this kind of attitudes, and I don't agree much with that. Immediately,

emotionally ... I have a bad feeling about that argument. [response category 4, disagree]

The fact that most Norwegian respondents 'recognize' certain items as belonging to a chauvinistic discourse, whereas such recognition is not seen in the Dutch sample, suggests that there is a difference in *salience* of items in Norway in comparison to the Netherlands. If this is the case, then items might not be equivalent between the two countries. Therefore, we conducted a separate study of the salience of the items of the IA Scale in the Norwegian and the Dutch population, in which this difference in salience could be confirmed (van der Veer *et al.*, 2003; data collected in February 2002). That study would not have taken place and its results would not have been obtained if the TSTI study reported in this paper had not discovered the existence of the phenomenon.

Conclusions

With regard to the IA Scale it can be concluded that the TSTI has identified the same (wording) problems that were identified with an expert review (not reported here), but also identified an additional problem. That problem concerns the way Dutch and Norwegian respondents interpret the items. The findings of our pilot study point to the possibility that specific kinds of ('literal') response behaviour might invalidate results obtained with the IA Scale. Subsequent research has shown that this discovery is very significant and has implications for the validity of the IA Scale.

An important conclusion about the TSTI is that not only is it able to confirm the wording problems identified with an expert review (not reported here), it also is able to detect other problems that are not discovered by expert review. TSTI results show in detail what different respondents actually do when they complete the IA Scale, and therefore it offers a more comprehensive diagnosis of the questionnaire as a whole. Results as reported in this article cannot be produced in a valid way with traditional 'cognitive' interviews because these do not focus on *observation* of actual response behaviour.

The main objective of this article is to demonstrate the usefulness of the TSTI in discovering problems with an attitude scale such as the IA scale. We do not claim that we have found all possible problems for all possible categories of respondents, which would require more heterogeneous (and much larger) samples than the convenience samples of undergraduate students used in this study. Because, as yet, the IA Scale has only actually been used in samples of undergraduate students, and because, therefore, published IA results pertain to student populations only, it was justified to evaluate the scale with a sample of such students.

If we compare the results of this study with our previous pilot of the TSTI on a set of questions about alcohol consumption (Jansen & Hak, 2005), we notice an important similarity and an important difference. The similarity is that in both cases the TSTI appears to be uniquely productive in identifying problems that arise from the (biographical, cultural, political) context in which the questionnaire is completed. The main difference between the two studies is with regard to the function of the third step of the TSTI (respondent validation). In the study of questions on alcohol consumption,

this third step provided much information on drinking behaviour that was valuable in explaining the difficulties that respondents encountered when they completed the questionnaire. In the study on attitudes towards illegal immigrants, relative little use is made of data collected in the third TSTI step. This difference can be explained by the fact that interviewers and respondents in the study on alcohol consumption questions used this third step for an exploration of the 'facts' of the respondents' drinking behaviour. In the study reported here it appeared that, in general, there was not much left to explore regarding the respondents' attitudes after they had completed the second step. The only issue that could be explored in the third step, which is reported above, was the respondent's attitude to the questionnaire (rather than to illegal immigrants).

We concluded from this application of the TSTI to an attitude scale that it is a useful tool for testing attitude questions. Together with our previous study, this confirms that the TSTI is a valuable addition to the repertoire of qualitative (pre-)testing techniques.

Note

- [1] This version of the IA Scale is the American version in which references to the USA have been substituted by Norway and the Netherlands and references to dollars by kroner, guilders and (more recently) euros. The term 'illegal alien' is American and should be replaced by a British equivalent if used in the UK (which has not happened as yet). Please note that respondents in our research filled out a Norwegian or Dutch version in which appropriate terminology regarding illegal immigrants was used.

References

- Bekerian, D. A., & Dennett, J. L. (1993). The cognitive interview technique: Reviving the issues. *Applied Cognitive Psychology*, 7, 275–297.
- Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis*. Cambridge, MA: MIT Press.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis* (rev. ed.). Cambridge, MA: MIT Press.
- Hak, T., van der Veer, K., & H. Jansen (2004). *The Three-Step Test-Interview (TSTI): An observational instrument for pretesting self-completion questionnaires*. ERIM Reports Series Research in Management. Retrieved from <https://ep.eur.nl/handle/1765/1265>
- Jansen, H., & Hak, T. (2005). The productivity of the three-step test-interview (TSTI) compared to an expert review of a self-administered questionnaire on alcohol consumption. *Journal of Official Statistics*, 21(1) pp. 103–120.
- Ommundsen, R., & Larsen, K. S. (1997). Attitudes toward illegal aliens: The reliability and validity of a Likert-type scale. *The Journal of Social Psychology*, 137(5), 665–667.
- Ommundsen, R., van der Veer, K., Mørch, S., Hak, T., & Larsen, K. S. (2002). Attitudes toward illegal immigration: A cross-national methodological comparison. *The Journal of Psychology*, 136(1), 103–110.
- van der Veer, K., Ommundsen, R., & Hak, T. (2004). Meaning shift of items in different language versions. A cross-national validation study of the Illegal Aliens Scale. *Quality & Quantity*, 37, 193–206.
- van der Veer, K., Ommundsen, R., Larsen, K. S., Van Le, H., Krumov, K., Pernice, R. E., & Romans, G. P. (2004). The structure of attitudes toward illegal immigration: The development of cross-national cumulative scales. *Psychological Reports*, 94, 897–906.
- Willis, G. B. (2005). *Cognitive interviewing. A tool for improving questionnaire design*. Thousand Oaks, CA: Sage.